Notice of Allowability	Application No.	Applicant(s)	
	10/613,366	WENG ET AL.	
	Examiner	Art Unit	
	Kandasamy Thangavelu	2123	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1. ☒ This communication is responsive to *July 9, 2007*.

2. ☒ The allowed claim(s) is/are *1-8 and 10-18*.

3. ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
   a) ☐ All b) ☐ Some* c) ☐ None of the:
      1. ☐ Certified copies of the priority documents have been received.
      2. ☐ Certified copies of the priority documents have been received in Application No. _____ .
      3. ☐ Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

   * Certified copies not received: _____.

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.
THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.

4. ☐ A SUBSTITUTE OATH OR DECLARATION must be submitted. Note the attached EXAMINER'S AMENDMENT or NOTICE OF INFORMAL PATENT APPLICATION (PTO-152) which gives reason(s) why the oath or declaration is deficient.

5. ☐ CORRECTED DRAWINGS (as "replacement sheets") must be submitted.
   (a) ☐ including changes required by the Notice of Draftsperson's Patent Drawing Review (PTO-948) attached
   1) ☐ hereto or 2) ☐ to Paper No./Mail Date _____.
   (b) ☐ including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date _____.

Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).

6. ☐ DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

Attachment(s)

1. ☐ Notice of References Cited (PTO-892)
2. ☐ Notice of Draftperson's Patent Drawing Review (PTO-948)
3. ☐ Information Disclosure Statements (PTO/SB/08), Paper No./Mail Date _____
4. ☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material
5. ☐ Notice of Informal Patent Application
6. ☐ Interview Summary (PTO-413), Paper No./Mail Date _____ .
7. ☒ Examiner's Amendment/Comment
8. ☒ Examiner's Statement of Reasons for Allowance
9. ☒ Other *Clean copy of allowed claims*.

DETAILED ACTION

Introduction

1. This communication is in response to the Applicants' communication dated July 9, 2007. Claims 1, 8, 10 and 12-18 were amended. Claims 1-18 of the application are pending.

Examiner's Amendment

2. Authorization for this examiner's amendment was given in a telephone conversation by Mr. Aaron Deditch on August 3, 2007.

An examiner's amendment to the record appears below. Should the changes and/or additions be unacceptable to the applicants, an amendment may be filed as provided by 37 CFR 1.312. To ensure consideration of such an amendment, it MUST be submitted no later than the payment of the issue fee.

3. In the claims:

1. A computer implemented method to select features for maximum entropy modeling for language and statistical processing, the method comprising:

(a) determining gains of log likelihood for candidate features during an initialization stage;

(b) ranking the candidate features in an ordered list based on the determined gains;

(c) selecting a top-ranked feature in the ordered list with a highest gain;

(d) adjusting a maximum entropy model using the selected top-ranked feature;

(e) determining gains of log likelihood for only a first predefined number of top-ranked features;

(f) repeating steps (b) through (e) until a number of selected features equals a second predefined number;

(g) storing the second predefined number of selected top-ranked features and the adjusted model in a file.

In Claim 6, Line 2, "pre-defined"

has been changed to

-- predefined --.

Replace claim 8 with:

8. A computer implemented method to select features for maximum entropy modeling for language and statistical processing, the method comprising:

(a) computing gains of log likelihood of candidate features using a uniform distribution;

(b) ordering the candidate features in an ordered list based on the computed gains;

(c) selecting a top-ranked feature with a highest gain in the ordered list;

(d) adjusting a maximum entropy model using the selected top-ranked feature;

(e) removing the top-ranked feature from the ordered list so that a next-ranked feature in the ordered list becomes the top-ranked feature and marking all features as not ranked;

(f) computing a gain of the top-ranked feature using the adjusted model;

(g) comparing the gain of the top-ranked feature with a gain of the next-ranked feature in the ordered list;

(h) if the gain of the top-ranked feature equals or is more than the gain of the next-ranked feature marking it as ranked and selecting the next-ranked feature that is not marked as ranked as the top-ranked feature;

(i) if the gain of the top-ranked feature is less than the gain of the next-ranked feature, repositioning the top-ranked feature in the ordered list so that the next-ranked feature becomes the top-ranked feature;

(j) repeating steps (f) through (i) until number of top-ranked features that are marked ranked equals a first predefined number;

(k) repeating steps (c) through (j) until one of a number of selected features equals a second predefined number and a gain of a last-selected feature falls below a predefined value; and

(l) storing the second predefined number of selected top-ranked features and the adjusted model in a file.

Cancel claim 9.

Replace claim 12 with:

12. A processing system to perform maximum entropy modeling in which one or more candidate features derived from a corpus of data are incorporated into a maximum entropy model that predicts linguistic behavior, the system comprising:

a computer with at least one processor, a memory storing a program of instructions and a display device;

a gain computation logic to determine gains of log likelihood for the candidate features during an initialization stage and to determine gains for only a first predefined number of top-ranked features during a feature selection stage;

a feature ranking logic to rank features based on the determined gains;

a feature selection logic to select a feature with a highest gain as a top-ranked feature; and

a model adjustment logic to adjust the maximum entropy model using the selected top-ranked feature;

wherein when the program is executed on the processor, a second predefined number of features with the highest gains are selected as the top-ranked features and included in the maximum entropy model; and

the second predefined number of selected top-ranked features and the adjusted model are stored in a file.

In Claim 13, Lines 1-2, "The hardware implemented processing arrangement system of claim 12, wherein feature ranking arrangement"

has been changed to

-- The processing system of claim 12, wherein feature ranking logic --.

In Claim 14, Lines 1-2, "The hardware implemented processing arrangement system of claim 12, wherein the gain computation arrangement"

has been changed to

-- The processing system of claim 12, wherein the gain computation logic --.

In Claim 15, Lines 1-2, "The hardware implemented processing arrangement system of claim 12, wherein the gain computation arrangement"

has been changed to

-- The processing system of claim 12, wherein the gain computation logic --.

In Claim 16, Line 1, "The hardware implemented processing arrangement system"

has been changed to

-- The processing system --.

In Claim 17, Line 1, "The hardware implemented processing arrangement system"

has been changed to

-- The processing system --.

Replace claim 18 with:

18. A computer storage medium having a set of instructions executable by a processor to perform maximum entropy modeling in which one or more candidate features derived from a corpus of data are incorporated into a maximum entropy model that predicts linguistic behavior comprising instructions for:

(a) ordering candidate features based on gains of log likelihood computed using a uniform distribution to form an ordered list of candidate features;

(b) selecting a top-ranked feature with a largest gain and adjusting the maximum entropy model for a next stage;

(c) removing the top-ranked feature from the ordered list of the candidate features so that a next-ranked feature in the ordered list becomes the top-ranked feature and marking all features as not ranked;

(d) computing a gain of the top-ranked feature using the adjusted model;

(e) comparing the gain of the top-ranked feature with a gain of the next-ranked feature in the ordered list;

(f) if the gain of the top-ranked feature equals or is more than the gain of the next-ranked feature marking it as ranked and selecting the next-ranked feature that is not marked as ranked as the top-ranked feature;

(g) if the gain of the top-ranked feature is less than the gain of the next-ranked feature, repositioning the top-ranked feature in the ordered list so that the next-ranked feature becomes the top-ranked feature;

(h) repeating steps (d) through (g) until number of top-ranked features that are marked ranked equals a first predefined number;

(i) repeating steps (b) through (h) until one of a number of selected features reaches a second predefined number and a gain of a last-selected feature falls below a predefined value; and

(j) storing the second predefined number of selected top-ranked features and the model in a file.

A clean copy of allowed claims is attached.

Reasons for Allowance

4. Claims 1-8 and 10-18 of the application are allowed over prior art of record.

5. The following is an Examiner's statement of reasons for the indication of allowable subject matter:

The closest prior art of record shows:

(1) a computerized language translation model for translating a series of source words into a series of target words; the model is a method of estimating the conditional probability that given x, the process will output y; the model is constructed using parameterized statistical modeling technique; a set of statistics that capture the predictive capacity of statistical models of natural language is used in a maximum entropy model; the model is adjusted by providing a set

of candidate features in the output process and using a score or gain representing the benefit of adding a feature to the model; the log likelihoods of all features are calculated and the feature with the maximum log likelihood is selected for addition to the model at each stage for the maximum entropy model; an incremental approach to selecting the features to be added to the model is applied (**Berger et al.,** U.S. Patent 6,304,841);

(2) a method and apparatus for determining the gain and training starting point of a feature function for maximum entropy/minimum divergence (MEMD) probability models in language modeling for speech recognition systems, language translation systems and grammar checking systems; an MEMD model is constructed from a base model and a set of feature functions from a given corpus of text; a large corpus will yield millions of potential features; to maximize the efficiency of the model, only those features that exhibit the highest predictive power is used in constructing the model; the MEMD modeling evaluates the gain of the candidate features, ranks them and retains those features that exhibit the highest gains; the method is applied to each of the candidate features to calculate their respective gain values; the features are ranked based on the computed gain values; features having an approximate gain value that exceeds a predetermined threshold value are output for use in constructing the MEMD model (**Printz,** U.S. Patent 6,049,767); and

(3) use of gain as a statistic and a figure of merit for selecting features for an MEMD language model; the model selects those features that that have highest predictive power for inclusion in the model; the method seeks features that improve upon the predictions of the training corpus; the gain is computed for the given features with respect to a base model; computing the gain for all the features in the corpus; features are ranked and a subset of features

selected for building the MEMD model; the method uses an iterative algorithm that selects one new feature for inclusion on each iteration (**Berger et al.,** "A comparison of criteria for Maximum entropy/minimum divergence feature selection", Submitted as part of the IDS).

None of these references taken either alone or in combination with the prior art of record discloses a computer implemented method to select features for maximum entropy modeling for language and statistical processing, specifically including:

(Claim 1) "(a) determining gains of log likelihood for candidate features during an initialization stage;

(b) ranking the candidate features in an ordered list based on the determined gains;

(c) selecting a top-ranked feature in the ordered list with a highest gain;

(d) adjusting a maximum entropy model using the selected top-ranked feature;

(e) determining gains of log likelihood for only a first predefined number of top-ranked features;

(f) repeating steps (b) through (e) until a number of selected features equals a second predefined number".

None of these references taken either alone or in combination with the prior art of record discloses a computer implemented method to select features for maximum entropy modeling for language and statistical processing, specifically including:

(Claim 8) "(c) selecting a top-ranked feature with a highest gain in the ordered list;

(d) adjusting a maximum entropy model using the selected top-ranked feature;

(e) removing the top-ranked feature from the ordered list so that a next-ranked feature in the ordered list becomes the top-ranked feature and marking all features as not ranked;

(f) computing a gain of the top-ranked feature using the adjusted model;

(g) comparing the gain of the top-ranked feature with a gain of the next-ranked feature in the ordered list;

(h) if the gain of the top-ranked feature equals or is more than the gain of the next-ranked feature marking it as ranked and selecting the next-ranked feature that is not marked as ranked as the top-ranked feature;

(i) if the gain of the top-ranked feature is less than the gain of the next-ranked feature, repositioning the top-ranked feature in the ordered list so that the next-ranked feature becomes the top-ranked feature;

(j) repeating steps (f) through (i) until number of top-ranked features that are marked ranked equals a first predefined number;

(k) repeating steps (c) through (j) until one of a number of selected features equals a second predefined number and a gain of a last-selected feature falls below a predefined value".

None of these references taken either alone or in combination with the prior art of record discloses a processing system to perform maximum entropy modeling, specifically including:

(Claim 12) "a gain computation logic to determine gains of log likelihood for the candidate features during an initialization stage and to determine gains for only a first predefined number of top-ranked features during a feature selection stage".

None of these references taken either alone or in combination with the prior art of record discloses a computer storage medium having a set of instructions executable by a processor to perform maximum entropy modeling, specifically including:

(Claim 18) "(b) selecting a top-ranked feature with a largest gain and adjusting the maximum entropy model for a next stage;

(c) removing the top-ranked feature from the ordered list of the candidate features so that a next-ranked feature in the ordered list becomes the top-ranked feature and marking all features as not ranked;

(d) computing a gain of the top-ranked feature using the adjusted model;

(e) comparing the gain of the top-ranked feature with a gain of the next-ranked feature in the ordered list;

(f) if the gain of the top-ranked feature equals or is more than the gain of the next-ranked feature marking it as ranked and selecting the next-ranked feature that is not marked as ranked as the top-ranked feature;

(g) if the gain of the top-ranked feature is less than the gain of the next-ranked feature, repositioning the top-ranked feature in the ordered list so that the next-ranked feature becomes the top-ranked feature;

(h) repeating steps (d) through (g) until number of top-ranked features that are marked ranked equals a first predefined number;

(i) repeating steps (b) through (h) until one of a number of selected features reaches a second predefined number and a gain of a last-selected feature falls below a predefined value".

6. Any comments considered necessary by applicant must be submitted no later than the payment of the issue fee and, to avoid processing delays, should preferably accompany the issue fee. Such submissions should be clearly labeled "Comments on Statement of Reasons for Allowance."

7. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Dr. Kandasamy Thangavelu whose telephone number is 571-272-3717. The examiner can normally be reached on Monday through Friday from 8:00 AM to 5:30 PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Paul Rodriguez, can be reached on 571-272-3753. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to TC 2100 Group receptionist: 571-272-2100.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only.

For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

K. Thangavelu
Art Unit 2123
August 3, 2007

Clean Copy of Allowed Claims

1. A computer implemented method to select features for maximum entropy modeling for language and statistical processing, the method comprising:

(a) determining gains of log likelihood for candidate features during an initialization stage;

(b) ranking the candidate features in an ordered list based on the determined gains;

(c) selecting a top-ranked feature in the ordered list with a highest gain;

(d) adjusting a maximum entropy model using the selected top-ranked feature;

(e) determining gains of log likelihood for only a first predefined number of top-ranked features;

(f) repeating steps (b) through (e) until a number of selected features equals a second predefined number;

(g) storing the second predefined number of selected top-ranked features and the adjusted model in a file.

2. The method of claim 1, wherein the gains of the candidate features determined in a previous feature selection stage are reused as upper bound gains of remaining candidate features in a current feature selection stage.

3. The method of claim 2, wherein the top-ranked feature is selected if its determined gain is greater than the upper bound gains of the remaining candidate features.

4. The method of claim 1, wherein the top-ranked feature is selected when a gain of the top-ranked feature determined using a currently adjusted model is greater than the gains of remaining candidate features determined using a previously adjusted model.

5. The method of claim 1, wherein gains for a predefined number of top-ranked features are determined at each feature selection stage.

6. The method of claim 1, further comprising:

re-evaluating gains of all remaining candidate features at a predefined feature selection stage.

7. The method of claim 1, wherein only the un-normalized conditional probabilities that satisfy a set of selected features are modified.

8. A computer implemented method to select features for maximum entropy modeling for language and statistical processing, the method comprising:

(a) computing gains of log likelihood of candidate features using a uniform distribution;

(b) ordering the candidate features in an ordered list based on the computed gains;

(c) selecting a top-ranked feature with a highest gain in the ordered list;

(d) adjusting a maximum entropy model using the selected top-ranked feature;

(e) removing the top-ranked feature from the ordered list so that a next-ranked feature in the ordered list becomes the top-ranked feature and marking all features as not ranked;

(f) computing a gain of the top-ranked feature using the adjusted model;

(g) comparing the gain of the top-ranked feature with a gain of the next-ranked feature in the ordered list;

(h) if the gain of the top-ranked feature equals or is more than the gain of the next-ranked feature marking it as ranked and selecting the next-ranked feature that is not marked as ranked as the top-ranked feature;

(i) if the gain of the top-ranked feature is less than the gain of the next-ranked feature, repositioning the top-ranked feature in the ordered list so that the next-ranked feature becomes the top-ranked feature;

(j) repeating steps (f) through (i) until number of top-ranked features that are marked ranked equals a first predefined number;

(k) repeating steps (c) through (j) until one of a number of selected features equals a second predefined number and a gain of a last-selected feature falls below a predefined value; and

(l) storing the second predefined number of selected top-ranked features and the adjusted model in a file.

9. Canceled.

10. The method of claim 8, wherein the gains of all remaining candidate features at a predefined feature selection stage are re-evaluated.

11. The method of claim 7, wherein gains of a majority of the candidate features remaining at each feature selection stage are reused based on a model adjusted in a previous feature selection stage.

12. A processing system to perform maximum entropy modeling in which one or more candidate features derived from a corpus of data are incorporated into a maximum entropy model that predicts linguistic behavior, the system comprising:

a computer with at least one processor, a memory storing a program of instructions and a display device;

a gain computation logic to determine gains of log likelihood for the candidate features during an initialization stage and to determine gains for only a first predefined number of top-ranked features during a feature selection stage;

a feature ranking logic to rank features based on the determined gains;

a feature selection logic to select a feature with a highest gain as a top-ranked feature; and

a model adjustment logic to adjust the maximum entropy model using the selected top-ranked feature;

wherein when the program is executed on the processor, a second predefined number of features with the highest gains are selected as the top-ranked features and included in the maximum entropy model; and

the second predefined number of selected top-ranked features and the adjusted model are stored in a file.

13. The processing system of claim 12, wherein feature ranking logic is configured to re-use gains of remaining candidate features determined in a previous feature selection stage using a previously adjusted model.

14. The processing system of claim 12, wherein the gain computation logic is configured to determine gains for top-ranked features in descending order from a highest to lowest until a top-ranked feature is encountered whose corresponding gain based on a current model is greater than gains of the remaining candidate features.

15. The processing system of claim 12, wherein the gain computation logic is configured to determine gains for a predefined number of top-ranked features at each feature selection stage.

16. The processing system of claim 15, wherein the predefined number of top-ranked features is 500.

17. The processing system of claim 12, wherein gains of all candidate features remaining at a predefined feature selection stage are re-evaluated.

18. A computer storage medium having a set of instructions executable by a processor to perform maximum entropy modeling in which one or more candidate features derived from a corpus of data are incorporated into a maximum entropy model that predicts linguistic behavior comprising instructions for:

(a) ordering candidate features based on gains of log likelihood computed using a uniform distribution to form an ordered list of candidate features;

(b) selecting a top-ranked feature with a largest gain and adjusting the maximum entropy model for a next stage;

(c) removing the top-ranked feature from the ordered list of the candidate features so that a next-ranked feature in the ordered list becomes the top-ranked feature and marking all features as not ranked;

(d) computing a gain of the top-ranked feature using the adjusted model;

(e) comparing the gain of the top-ranked feature with a gain of the next-ranked feature in the ordered list;

(f) if the gain of the top-ranked feature equals or is more than the gain of the next-ranked feature marking it as ranked and selecting the next-ranked feature that is not marked as ranked as the top-ranked feature;

(g) if the gain of the top-ranked feature is less than the gain of the next-ranked feature, repositioning the top-ranked feature in the ordered list so that the next-ranked feature becomes the top-ranked feature;

(h) repeating steps (d) through (g) until number of top-ranked features that are marked ranked equals a first predefined number;

(i) repeating steps (b) through (h) until one of a number of selected features reaches a second predefined number and a gain of a last-selected feature falls below a predefined value; and

(j) storing the second predefined number of selected top-ranked features and the model in a file.